

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification <sup>6</sup> :	A1	(11) International Publication Number: <b>WO 99/12031</b>
G01N 33/483, C12Q 1/68		(43) International Publication Date: 11 March 1999 (11.03.99)

(21) International Application Number: PCT/AU98/00700	(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, HR, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).
(22) International Filing Date: 28 August 1998 (28.08.98)	
(30) Priority Data: PO 8824 29 August 1997 (29.08.97) AU	
(71) Applicants (for all designated States except US): SOUTHERN CROSS UNIVERSITY [AU/AU]; Military Road, East Lismore, NSW 2480 (AU). GRAINS RESEARCH & DEVELOPMENT CORPORATION [AU/AU]; Level 2, Bligh House, National Circuit, Barton, ACT 2600 (AU).	
(72) Inventors; and	
(75) Inventors/Applicants (for US only): HENRY, Robert, James [AU/AU]; 89 Bundara Park Drive, Tuckombil, NSW 2477 (AU). SHEPHERD, Mervyn [AU/AU]; 17A Frederick Street, Lismore, NSW 2480 (AU).	
(74) Agent: KELLY, Robin, Thomas; Fisher Adams Kelly, Level 13, AMP Place, 10 Eagle Street, Brisbane, QLD 4000 (AU).	

(54) Title: A METHOD OF GENOTYPING

(57) Abstract

A method of genotyping is provided with wide applicability to all organisms including plants, animals, bacteria and viruses. The method essentially relies on detecting melt curve or  $T_m$  differences between DNA samples derived from organisms for which genotypic analysis is required. An advantage of this method is that a single  $T_m$  value may be obtained from DNA samples containing multiple DNA fragments, thereby allowing an accurate determination of genetic similarity or dissimilarity between samples.

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bohnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

TITLE**"A METHOD OF GENOTYPING"**FIELD OF THE INVENTION

THIS INVENTION relates to a method for determining  
5 genetic relationships between individuals and groups of individuals.

BACKGROUND OF THE INVENTION

In relation to the background of the invention, it is intended  
to split this into two parts, i.e. (i) conventional genotyping techniques and  
(ii) detection and comparison of DNA sequences based on melting  
10 temperature.

(i) Conventional Genotyping Techniques

"Genotyping", as used hereinafter, is a process for  
determining genetic relatedness between individuals and groups of  
individuals, according to genotype. The genetic makeup characteristic of  
15 a particular individual or group of individuals will hereinafter be referred to  
as a "genotype". For the purposes of this specification, an "individual" is  
any biological entity possessing a genome, and can include an acellular  
entity such as a virus. A "group of individuals" is two or more such  
individuals.

20 Genotype is often characterized according to regions of  
genomic DNA that are "polymorphic" with respect to nucleotide sequence.  
"Polymorphism" is a feature characteristic of genomes, and refers to the  
occurrence of variable forms of one or more regions of DNA therein.  
According to this definition, "polymorphic" regions of DNA vary between  
25 individuals or groups of individuals in terms of sequence length, base  
composition, continuity or degree of repetition. "A polymorphism" is  
therefore a variant of a particular region of genomic DNA, for example.  
Such polymorphisms can be inherited, can arise somatically, or through  
specialized mechanisms such as immunoglobulin gene rearrangement  
30 and hypermutation. Different individuals, or groups of individuals, can  
therefore be distinguished genotypically according to these

**polymorphisms**

Genotyping is widely used in plant and animal breeding, forensic science, forestry, the study of inherited disease, organ donation and transplantation. The importance of genotyping resides in the fact a 5 particular genotype often underlies a phenotypic trait. For example, in order to introduce a phenotypic trait into an individual or group of individuals, such as through selective breeding, it is important to be able to correlate a particular phenotypic trait with a particular genotype.

In some cases, a polymorphic region of DNA useful for 10 genotyping is a "locus". Usually, a locus correlates with a specific phenotypic trait, and in many cases encodes a protein. Alternative forms of a particular polymorphic locus are designated "alleles". A typical example in humans would be the various heritable allelic forms of the major histocompatibility complex loci. However, other polymorphisms are 15 also useful in genotyping, in particular, those that include repetitive DNA elements such as minisatellites (Jeffreys *et al.*, 1985, *Nature* **314** 67), microsatellites (Tautz & Renz, 1984, *Nucl. Acid. Res.* **12** 4127) and variable number tandem repeats (Nakamura *et al.*, 1987, *Science* **235** 1616). Such repetitive sequences are rapidly evolving (Flavell *et al.*, 20 1982, *Chromosomal DNA sequences and their organisation. In: Encyclopaedia of Plant Physiology, New Series, 14B* 46-74) and highly polymorphic between individuals and groups of individuals (Tautz *et al.*, 1986, *Nature* **322** 652), and are therefore very suitable for genotyping.

One of the first methods used for genotyping is restriction 25 fragment length polymorphism (RFLP) analysis (Nathans & Smith, 1975, *Ann. Rev. Biochem.* **46** 273). This method takes advantage of polymorphisms that result in the loss or appearance of restriction enzyme sites in polymorphic regions of DNA. The presence or absence of these 30 restriction sites results in differences in the size of DNA fragments resulting from restriction enzyme digestion. Size fractionation of the digested DNA followed by hybridization with a probe complementary to

the expected restriction fragment allows the identification of differences in restriction fragment length, indicative of polymorphic variation.

More recently, methods employing nucleic acid amplification techniques, such as the polymerase chain reaction (PCR), have become important in genotyping. Such techniques allow:-

- (i) the detection of nucleotide changes that create mismatched PCR primer annealing sites and thus differences in PCR product intensity or the presence or absence of PCR product;
- 10 (ii) the detection of "lost" primer annealing sites, such as through deletion events or hypermutation, by observing the presence or absence of a PCR product; and
- 15 (iii) the detection of differences occurring through insertion or deletion of intervening nucleotide sequences, by comparison of PCR product size.

All of the abovementioned methods require knowledge of sufficient nucleotide sequence to allow analysis of specific polymorphic loci. However, alternative PCR-based genotyping methods have been devised that do not require such knowledge. Randomly amplified polymorphic DNA (RAPD) sequences are PCR products generated using primers of arbitrary sequence which anneal at multiple sites within a given DNA template (Williams *et al.*, 1990, *Nucl. Acid. Res.* **18** 6531; Welsh & McClelland, 1990, *Nucl. Acid. Res.* **18** 7213). The resultant pool of PCR products constitute a "fingerprint" characteristic of a particular genotype, without the identity of any of the products necessarily being known.

Amplified fragment length polymorphism (AFLP) can also be performed without prior sequence knowledge (Vos *et al.*, 1995, *Nucl. Acid. Res.* **23** 4407). This technique, like RFLP, takes advantage of restriction site variability, and uses a PCR approach similar to RAPD in order to amplify the restriction fragments to the required level of

detectability. Again, the amplified fragments provide a "fingerprint" characteristic of a particular genotype.

A characteristic of polynucleotide sequences that include regions of complementary base-pairing, is that at a particular 5 temperature, hereinafter referred to as the "melting temperature" or " $T_m$ ", complementary base-pairing is disrupted. The characteristic  $T_m$  of a particular polynucleotide sequence depends primarily upon the guanine (G) and cytosine (C) base content of the sequence, and to a lesser extent on the length of the polynucleotide sequence. A commonly used empirical 10 formula describing this relationship is:-

$$T_m = K + 0.41(\%G+C) - 500/L$$

where  $K = \log_{10} \{[\text{salt}]/1.0 + 0.7x[\text{salt}]\}$

and  $L$  = sequence length

and  $[\text{salt}]$  = the molar concentration of salt present in the 15 preparation that contains the polynucleotide sequence; salt would include the chloride or acetate salts of sodium, potassium, magnesium or ammonium, such as commonly used in preparations of polynucleotide sequences.

Additionally, G + C distribution within a polynucleotide 20 sequence can also affect  $T_m$ : GC "clamps" at each end of a polynucleotide sequence can increase  $T_m$  without increasing the overall G + C content.

Thus, differential  $T_m$  can be used as an indicator of nucleotide sequence differences due to sequence length, base 25 composition and base distribution. A method used in genotyping that takes advantage of  $T_m$  is Denaturing Gradient Gel Electrophoresis (DGGE; Folde & Loskoot, 1994, Hum. Mut. 3 83). According to this method, DNA fragments are electrophoresed through a gel made with a gradient of denaturant such as urea, so that as the DNA fragment 30 encounters increasing concentrations of denaturant, the  $T_m$  is effectively lowered due to disruption of complementary base-pairing by the

denaturant. The strands therefore begin to dissociate, resulting in a more "relaxed" conformation and impaired migration through the gel.  $T_m$  differences between DNA fragments derived from specific polymorphic loci can therefore be detected on the basis of gel migration differences.

5 All of the above methods use gel electrophoresis, and DNA fragments are visualized either by ethidium bromide staining or by hybridization with specific probes. Such probes can be prepared by radiolabeling, or adapted for chromogenic detection by the incorporation of enzymes such as alkaline phosphatase, or for chemiluminescent detection via horseradish peroxidase. The chromogenic detection method 10 has been incorporated into microscale detection systems such as those based on the ELISA method.

15 Major disadvantages, therefore, are that gel electrophoresis together with techniques for visualization of DNA fragments require multiple complex steps, are time consuming, potentially hazardous and difficult to automate for the purposes of high-throughput genotyping.

(ii) Detection and comparison of DNA sequences based on melting temperature

20 Methods of analyzing polynucleotide sequences according to melting characteristics have been developed. These methods can be categorized into three different types:-

- (1) measuring a signal from a specifically bound reporter molecule which indicates the presence or absence of a polynucleotide sequence;
- 25 (2) generating a "melt curve" as an indicator of the presence or absence of a polynucleotide sequence; and
- (3) an extended version of (2), whereby melt curves are compared, each melt curve being characteristic of a distinct polynucleotide sequence.

30 With regard to (1), this type of method utilizes reporter

molecules, and in particular, fluorescent reporter molecules, which bind to polynucleotide sequences. In a typical example of this method, temperature is increased up to and beyond the  $T_m$  so that base-pairing is disrupted. If the fluorescent reporter molecule is irradiated with incident 5 light of a suitable wavelength to excite light emission by the fluorescent molecule, the disruption of base-pairing that occurs as temperature increases would be indicated by measuring a decrease in a light emission signal provided by the fluorescent molecule.

This decrease in emission is primarily due to the fact that 10 the amount of light emitted per fluorescent molecule is much greater when the molecule is bound to polynucleotide sequences that contain substantial regions of complementary base-pairing, such as double-stranded DNA (or to a lesser extent single-stranded DNA that includes regions of internal base-pairing). In the case of the fluorescent molecule 15 EtBr, intercalation of the molecule between the hydrophobic faces of adjacent base pairs greatly enhances light emission upon excitation with UV light, so that disruption of base-pairing results in a marked decrease in the amount of light emitted per bound molecule (Higuchi *et al.*, 1992, *Biotechnology* **10** 413).

20 Use of fluorescence emission by EtBr for determining the presence of PCR product by Higuchi *et al.*, 1992, *supra* was subsequently applied to the real-time monitoring of PCR product formation (Higuchi *et al.*, 1993, *Biotechnology* **11** 1026). In this latter method, (EtBr) was added to PCR reactions to provide a fluorescence signal suitable for real-time 25 monitoring of the formation of DNA product.

With regard to (2), the first application of melt curve analysis 30 was restricted to detecting the "end-point" formation of a PCR product, as disclosed in JP 3-147796A. According to this reference, the presence of a specific double-stranded PCR product could be distinguished from "background" polynucleotide sequences. This involved increasing the temperature while measuring either:-

- (i) a fluorescence signal emitted by DNA-bound Ethidium Bromide (EtBr); or
- (ii) measuring an absorbance signal in the absence of EtBr;

5 as the PCR product conformation changed from double-stranded to single stranded. The absorbance signal resulted from the PCR product DNA absorbing a portion of incident light, usually 260 nm light, the absorbance being measured by a spectrophotometer. The absorbance of 260 nm light by single stranded DNA is greater than that by double-stranded DNA.

10 This reference therefore defined two types of "melt curve" :-

- (i) a Fluorescence ( $F$ ) vs Temperature ( $T$ ) melt curve, constructed by plotting a fluorescence signal, measured from the EtBr reporter molecule added to the DNA samples, against temperature; and
- (ii) an Absorbance ( $A$ ) vs Temperature ( $T$ ) melt curve constructed by plotting absorbance measurements, in the absence of a reporter molecule, against temperature.

20 All of the methods applicable to (1) and (2) were essentially restricted to determining the presence or absence of a specific PCR product, or providing a semi-quantitative indication of the amount of PCR product formed.

25 The first use of melt curves as a means of comparing polynucleotide sequences, as in (3), is provided in European patent application EP 711840. In this reference, the DNA analysed was in single-stranded form and the method was limited to single-stranded sequences capable of substantial internal base-pairing. According to this reference, single-stranded DNA "samples" were derived from PCR products amplified from a HLA-DQA1 locus of unknown haplotype. The melt curves obtained from these single-stranded DNA samples of unknown HLA-DQ1 30 haplotype were then compared with melt curves obtained from single-

stranded DNA molecules corresponding to each known HLA-DQ1 haplotype, in order to identify the haplotype of each sample. This reference also contemplated the use of either  $A$  vs  $T$  or  $F$  vs  $T$  melt curves as disclosed in JP 3-147796.

5 Reference may also be made to Ririe *et al.*, 1997, *Anal. Biochem.* **245** 157 in which it was shown that distinct double-stranded DNA PCR products could be distinguished not only by comparing melt curves, but also by comparing data derived therefrom, such as  $T_m$  values. A microvolume fluorimeter integrated with a thermal cycler (*Light Cycler*, 10 Idaho Technology, USA) was used to monitor PCR product formation either during amplification, or as an "end-point" product, by measuring light emitted by the fluorescent reporter molecule SYBR Green I, which displays greater selectivity for DNA and more enhanced fluorescence when bound to double-stranded DNA compared to EtBr.

15 By comparing the amount of fluorescence emitted ( $F$ ) versus temperature ( $T$ ), Ririe *et al.*, 1997, *supra* created  $F$  vs  $T$  melt curves, characteristic of particular PCR products. Furthermore, melt curves could be mathematically analyzed to provide the negative derivative of the curve function, which could be plotted as a  $-dF/dT$  vs  $T$  melt curve to 20 provide:-

- (i) a peak from which a  $T_m$  value could be derived; and
- (ii) a peak which could be integrated to provide an upper limit to the ratio of specific product to total product.

25 This method provided rapid analysis of samples in as little as 15 minutes, which far surpasses electrophoretic and hybridization methods which routinely take hours or days.

30 As a result, Ririe *et al.*, 1997, *supra* could discriminate between the PCR fragments amplified from markedly different loci, for example a 536 bp fragment of the human  $\beta$  globin gene and a 180 bp fragment of the Hepatitis B virus surface antigen gene, by visual comparison of either  $F$  vs  $T$  melt curves or  $-dF/dT$  vs  $T$  melt curves, or

according to  $T_m$  values derived therefrom. Furthermore, mixtures of the two products provided a composite  $F$  vs  $T$  melt curve which nevertheless could be resolved into distinct  $-dF/dT$  vs  $T$  melt curves and distinct  $T_m$  values (Ririe *et al.*, 1997, *supra*).

5        This method was predicted to be capable of discriminating between very similar polynucleotide sequences indicative of genotype. A simple example proposed by Ririe *et al.*, 1997, *supra*, was in the case where a heterozygote possesses two allelic forms of a single locus. DNA fragments which were PCR amplified from the heterozygotic locus would  
10      comprise a mixed population of heteroduplexes and homoduplexes as a result of each cycle of denaturation and re-annealing during PCR. In this regard, heteroduplex DNA comprises single strands with partially complementary nucleotide sequences due to limited differences in nucleotide sequence or "base-pair mismatches", which strands anneal to  
15      form double-stranded DNA notwithstanding base-pair mismatches. It was proposed that in principle, it should be possible to distinguish the heteroduplex DNA from the relevant homoduplex DNA by virtue of their exhibiting different melt curves, the heteroduplex having a lower  $T_m$  than the homoduplex.

20        In a subsequent confirmation of this prediction, it was shown using melt curve analysis that a heteroduplex with a three base-pair mismatch could be distinguished from the relevant homoduplex DNA (*LightCycler: Methods and Operations Manual*, Idaho Technology).

25        In this regard, reference is also made to International Publications WO97/46707, WO97/46712 and WO97/46714 which disclose apparatus and methods applicable to fluorescent monitoring of DNA amplification, such as provided by the *LightCycler*. These publications disclose both real-time and end-point monitoring of PCR reactions, quantitation of PCR products by melt curve analysis, and the  
30      use of fluorescently-labeled hydrolysis probes, fluorescence energy resonance transfer (FRET) probes, and SYBR Green I and EtBr in melt

curve analysis.

In WO97/46707, the usefulness of reporter molecules such as SYBR Green I in genotyping based on discrimination between homozygotes and heterozygotes is reiterated (as originally reported in 5 Ririe *et al.*, 1997, *supra*). An example is provided where a common cystic fibrosis mutation comprising a 3 base pair deletion is detected by melt curve analysis. However, an admitted intrinsic limitation of this method is that it cannot distinguish between homozygous wild-type and homozygous mutant individuals.

10 Thus, WO97/46714 has proposed the superiority of sequence-specific fluorescence energy transfer (FRET) hybridization probes for genotyping applications. This technique relies on the use of a labeled probe with a specific sequence which differentially hybridizes to wild-type and mutant allelic sequences, differential hybridization resulting 15 from base-pair mismatches with either the mutant or wild-type sequence. The exquisite sensitivity of this method in discriminating between similar DNA molecules allowed the detection of single-base differences between allelic forms of single genetic loci. Unlike melt curve analysis based on discrimination between heteroduplexes and homoduplexes, this method 20 allows discrimination between heterozygote, homozygote wild type and homozygote mutant individuals. This was shown for the factor V Leiden mutation (a single G to A base change) and a common point mutation in the human methylenetetrahydrofolate (MTHFR) gene.

In summary, genotyping by melt curve analysis of double-stranded DNA has proceeded along two lines:-

- (i) identification of allelic differences between heterozygotes and homozygotes by detection of heteroduplex formation; and
- (ii) detection of known mutations through base-pair mismatches between a fluorophore-labeled hybridization probe and either a wild-type or mutant

allelic sequence resulting in  $T_m$  differences.

In both cases, these methods have fundamentally relied on detecting melt curve differences or similarities resulting from the respective presence or absence of base-pair mismatches. Thus, prior art 5 genotyping has focussed exclusively on detecting melt curve and  $T_m$  differences between only those polynucleotide sequences which represent known allelic forms of a single locus.

#### OBJECT OF THE INVENTION

The present inventors have sought to extend the range of 10 polynucleotide sequences available to genotyping by melt curve analysis. In many cases, methods which produce polynucleotide sequences useful in genotyping actually produce a multiplicity of polynucleotide sequences. Furthermore, the nucleotide sequence composition of such polynucleotide 15 sequences, and any differences therebetween, may not be known or predictable in advance.

Based on the prior art, it would be expected that melt curve comparisons of samples containing a variety of polynucleotide sequences would be extremely difficult, because of the likelihood that multiple  $-dF/dT$  vs  $T$  melt curves and  $T_m$  values would be obtained from such samples. In 20 particular, it would be expected that resolution of statistically significant differences between samples having multiple similar  $-dF/dT$  vs  $T$  melt curves and  $T_m$  values would be overly complicated and potentially inaccurate.

Surprisingly, the present inventors have found that multiple 25 polynucleotide sequences do not necessarily produce correspondingly complex  $-dF/dT$  vs  $T$  melt curves or multiple  $T_m$  values derived therefrom, and that melt curves or  $T_m$  values so derived may be utilized to compare 30 genotypes. In contrast to what might have been expected from the prior art, the present inventors have found that each polynucleotide sequence in a sample does not necessarily produce a distinct  $-dF/dT$  vs  $T$  melt curve or  $T_m$  value derived therefrom.

It is therefore an object of the present invention to provide a method of comparing individuals and/or groups of individuals according to genotype, which method does not suffer from the limited applicability of prior art methods.

5

#### SUMMARY OF THE INVENTION

The present invention therefore resides in a method of comparing genotypes, said method including the steps of:-

- (i) producing two or more samples, each of which includes two or more double-stranded polynucleotide sequences representative of a genotype of an individual or group of individuals;
- (ii) adding a fluorescent reporter molecule to the samples produced in step (i);
- (iii) modulating temperature whereby a conformation of said two or more double-stranded polynucleotide sequences is altered, and measuring a signal emanating from said reporter molecule, which signal changes as the conformation of said one or more double-stranded polynucleotide sequences is altered;
- (iv) using the signal measured in (iii) to construct a single  $-dF/dT$  melt curve corresponding to the two or more polynucleotide sequences; and
- (v) comparing respective single  $-dF/dT$  melt curves, or single  $T_m$  values derived from each melt curve, in order to identify a genetic relationship between each said individual or group of individuals.

25

The method of the invention is therefore suitable for comparing the genotype of any individuals or groups of individuals to assess the genetic relationship therebetween. Preferably, these would be cellular biological entities, such as animals, plants and bacteria, but most preferably would be cereal plants. The term "group" in this context

30

encompasses any collection of two or more individuals. Typical examples of such groups apparent to the skilled addressee would include strain, cultivar, breed, race, species, genus, family, order, class, phylum, kingdom and subgroups thereof.

5 In this regard, "a genetic relationship" refers to any kind of relationship that can be determined to exist between said individuals or groups of individuals compared by the method of the invention, whether the relationship be one of genetic identity, similarity, or dissimilarity.

10 The term "polynucleotide sequence" as used herein designates RNA, cDNA or DNA. Double-stranded polynucleotide sequences could therefore comprise pairs of RNA, DNA or cDNA single strands in any combination. A skilled addressee will appreciate that cDNA is complementary DNA produced from an RNA template by the action of reverse transcriptase. Furthermore, if the RNA template has been 15 processed to remove introns, the cDNA will not be identical to the gene from which the RNA was transcribed.

20 Preferably, said two or more double-stranded polynucleotide sequences would be produced by PCR amplification. Other suitable polynucleotide sequence amplification techniques well known to the skilled addressee, such as strand displacement amplification (SDA) and rolling circle replication (RCR), could also be applied. Alternatively, said 25 two or more polynucleotide sequences could be obtained by purifying fragments of restriction enzyme digested genomic DNA, cDNA, plasmid, bacteriophage, phagemid, cosmid or yeast artificial chromosome vectors containing said one or more polynucleotide sequences. This approach could include generating an array of polynucleotide fragments by restriction-enzyme digestion, as in RFLP analysis (Nathans & Smith, 1975, Ann. Rev. Biochem. **46** 273 hereinafter incorporated by reference).

30 Thus, said two or more polynucleotide sequences may be "representative of a genotype" by corresponding to regions of DNA, or RNA. In this regard, it will be understood by a skilled person that said two

or more polynucleotide sequences could correspond to one or more regions of the genome of an RNA virus.

Preferably, said two or more polynucleotide sequences correspond to one or more regions of genomic DNA; these might be 5 regions of genomic DNA selected on the basis of known polymorphism, or regions of genomic DNA not selected as such. Regions of DNA not selected on the basis of known polymorphism could be prepared by amplification of a template using techniques such as RAPD (Williams *et al.*, 1990, *Nucl. Acid. Res.* 18 6531 and Welsh & McClelland, 1990, *Nucl. Acid. Res.* 18 7213; herein incorporated by reference) and AFLP (Vos *et al.*, 1995, *Nucl. Acid. Res.* 23 4407; herein incorporated by reference) 10 which randomly amplify a plurality of polynucleotide sequences to create a "fingerprint" useful in genotyping.

In many cases, techniques such as RFLP, RAPD, AFLP and 15 techniques which amplify microsatellite and ribosomal repeat sequences generate two or more polynucleotide sequences, and are therefore particularly applicable to the invention. An example of genotyping using amplified microsatellite (wheat WMS44) and ribosomal (Rrn5 locus of cereals) repeat sequences, together with bush rat samples generated by 20 RAPD, is provided hereinafter.

It should also be noted with the aforementioned in mind, that "two or more polynucleotide sequences" includes, for example, two or more allelic forms of a locus, even where a nucleotide sequence difference between each allele is a single base.

25 Suitably, the signal changes in response to an alteration in conformation of said two or more polynucleotide sequences. In this regard, conformation refers to the aspect of three-dimensional structure resulting from complementary base-pairing. Complementary bases of single-stranded polynucleotide sequences pair to produce double-stranded polynucleotide sequences, or pair to produce internally base-paired single stranded polynucleotide sequences according to base- 30

pairing rules.

In DNA, the complementary bases are:-

- (i) A and T
- (ii) C and G.

5

In RNA, the complementary bases are:

- (i) A and U
- (ii) C and G.

It will be understood by the skilled person that under conditions favouring complementary base-pairing, polynucleotide 10 sequences will be substantially double-stranded. Under conditions that prevent or disrupt complementary base-pairing, polynucleotide sequences will be substantially single-stranded and lacking in internal base-pairing.

Accordingly, as used herein "modulating temperature 15 whereby a conformation of said two or more double-stranded polynucleotide sequences is altered" refers to modulating temperature such that complementary base-pairing is disrupted. Temperature modulation could be performed either by starting from a temperature at which base-pairing is substantially prevented or disrupted, preferably 20 95°C, and decreasing said temperature preferably to 60°C, or by starting at a temperature at which base-pairing is substantially favoured, preferably 60°C, and increasing said temperature, preferably to 95°C, to substantially disrupt base-pairing. Most preferably, said temperature would be increased from 60°C to 95°C.

25

Suitably, the fluorescent reporter molecule is capable of binding said two or more polynucleotide sequences and capable of providing a fluorescence signal which changes in response to an alteration in conformation of said two or more polynucleotide sequences.

30

Preferably, the intensity of fluorescence signal emanating from the fluorescent reporter molecule diminishes in magnitude as a double-stranded polynucleotide sequence is altered to a substantially

single-stranded conformation.

Suitably, the fluorescent reporter molecule is selected from the group consisting of Chromomycin A3, Pico Green, SYBR Green I, Ethidium Bromide, Acridine Orange, Thiazole Orange, and YO-PRO-1. A 5 discussion of the physico-chemical properties of many of these reporter molecules and their relative advantages and disadvantages is provided in Fu *et al.*, 1991, Cell 67 1047, which is herein incorporated by reference.

It will be appreciated SYBR Green I or EtBr are preferred reporter molecules.

10 Most preferably, SYBR Green I is the reporter molecule applicable to the method of the invention. SYBR Green I binds DNA and provides a fluorescence signal whereby the magnitude of fluorescence emitted per molecule of SYBR Green I is greatest when the reporter molecule is bound to double-stranded DNA.

15 Measurement of said signal provided by the fluorescent reporter molecule could be performed by a device such as a fluorimeter, or by a flow cytometer. Preferably, detection would be performed by a fluorimeter capable of modulating temperature and handling multiple samples simultaneously. For example, detection may be performed using 20 the *LightCycler* apparatus manufactured and sold by Idaho Technology, ID USA.

25 In such a case, increasing temperature leads to the disruption of base-pairing and thus alter the conformation of said one or more polynucleotide sequences from double-stranded to substantially single-stranded, and result in a decrease in the magnitude of fluorescent signal provided by the DNA-bound fluorescent reporter molecule.

30 Initially, signal measurements may be used to first construct an  $F$  vs  $T$  melt curve, from which is derived a  $-dF/dT$  vs  $T$  melt curve such as shown in FIG. 2 and FIG. 4. Preferably, the  $-dF/dT$  melt curve has a single resolvable peak which is used to establish a single  $T_m$  value as shown in TABLE 1, TABLE 3 and TABLE 4. In some cases, a  $-dF/dT$  melt

curve may display two resolvable peaks. In such cases, the peak occurring at a higher temperature rather than a lower temperature is used to derive a single  $T_m$  value. The rationale is two-fold: (i) the higher temperature corresponds to the temperature at which substantially all of 5 the polynucleotide sequences have melted; and (ii) experimental results described hereinafter have suggested that the majority of polynucleotide sequences melt at the higher temperature rather than the lower temperature. Thus a single  $T_m$  value derived from the higher temperature peak would appear to be more generally representative of the 10 polynucleotide sequences in the sample.

Comparison of respective  $-dF/dT$  vs  $T$  melt curves can also be performed visually. However, particularly in cases where there are small differences between curves, derivation of  $T_m$  values and comparison thereof is the preferred means whereby melt curves can be accurately 15 compared. The generation of  $F$  vs  $T$  and  $-dF/dT$  vs  $T$  melt curves,  $T_m$  values, and comparisons therebetween are performed as hereinafter described.

Preferably, the graphical, mathematical and statistical operations applicable to the method of the invention are performed by 20 utilizing a computer together with appropriate software. An example of preferred methods of statistical analysis and computer software applicable thereto is provided hereinafter.

It will be appreciated that the method of the invention has a number of possible applications. It could be used to compare the 25 genotypes of two or more individuals or groups of individuals simply to determine whether they are identical, distinct or perhaps related, without regard to establishing identity. A more refined application would be where melt curves derived from one or more unknown individuals or groups of individuals are compared with melt curve "standards" representative of the genotypes of known individuals or groups of individuals. This latter 30 application would allow a skilled addressee to establish the identity of the

unknown individual or group of individuals.

#### BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1: Agarose gel electrophoresis of samples containing PCR products amplified from cereal Rrn5 ribosomal repeat locus.

5 FIG. 2: Representative  $-dF/dT$  vs  $T$  melt curves for each sample in FIG. 1.

FIG. 3: Agarose gel electrophoresis of samples containing PCR products amplified from wheat WMS 44 microsatellite locus.

10 FIG. 4: Representative  $-dF/dT$  vs  $T$  melt curves for each sample in FIG. 3.

FIG. 5: Agarose gel electrophoresis of samples containing RAPD fragments amplified from genomic DNA of individual bush rats.

## EXPERIMENTAL

### MATERIALS AND METHODS

#### **1.1 PCR amplification of 5S Ribosomal sequences**

##### **1.1.1 Template**

20 Genomic DNA preparations from rye, rice, wheat, maize oats and barley were obtained using a modified CTAB method (Graham et al., 1994, Biotechnology 16 48). 100 ng of template was used per PCR reaction.

##### **1.1.2 Primers**

25 Non-species specific consensus primers for the Rrn5 gene were synthesized according to SEQ ID NO:1 (forward) and SEQ ID NO:2 (reverse).

Each primer was used at 0.55  $\mu$ M per PCR reaction.

##### **1.1.3 Thermal Cycling**

30 PCR was performed using a *LightCycler* (Idaho Technology, Idaho USA) using the following conditions: 15 seconds at 94°C followed by 35 cycles each comprising a ramp to 94°C for 0 seconds, ramp to 56°C

for 0 seconds and then a ramp to 72°C for 30 seconds; these 35 cycles were then followed by a hold at 72°C for 30 seconds. Each PCR reaction comprised a 10 µl volume including 50 mM Tris-HCl pH 8.3, 2 mM MgCl<sub>2</sub>, 0.5 mg/ml bovine serum albumin, 200 µM each dNTP and 0.4 Units Taq polymerase. Each reaction also included SYBR Green I (Molecular Probes) at a 1:50,000 of the concentration provided by the manufacturer.

5 1.2. PCR product analysis

10 10 µl of each reaction was analysed on a 2.0% agarose gel stained with 0.5 µg/ml EtBr and PCR products visualized under UV light.

10 Each reaction was subjected to melt curve analysis as described below.

15 2.1 PCR amplification of WMS44 microsatellite sequences

2.1.1 Template

15 Genomic DNA was obtained from the seedling tissue of 5 common wheat varieties: *Gamenya*, *Halbert*, *Cunningham*, *Molineux*, and *Chinese Spring*, and two wheat/rye hybrids *Tiga* and *Tahara*, using the CTAB method as above. 100 ng of each template was used per PCR reaction.

20 2.1.2 Primers

20 Primers based on sequences flanking the wheat microsatellite region WMS44 (Roder et al., 1995, Mol. Gen. Genet. **246** 327-333), this region defined by the repeat sequence (GA)<sub>20</sub>, were synthesized according to SEQ ID NO:3 (forward) and SEQ ID NO:4 (reverse).

25 2.1.3 Thermal cycling

25 PCR was performed using a PE 9600 Gene Amp Thermal Cycler (Perkin Elmer, Foster City, CA, USA), using the following conditions: 35 cycles of 96°C for 1 minute, 60°C for 1 minute and 72°C for 2 minutes, after which samples were held at 72°C for 10 minutes. Each PCR reaction comprised a 25 µl volume including 10 mM Tris-HCl, 1.5 mM MgCl<sub>2</sub>, 50 mM KCl, 200 µM each dNTP and 0.4 Units Taq polymerase.

2.2. PCR product analysis

10  $\mu$ l of each PCR reaction product was analyzed by 10% polyacrylamide gel electrophoresis, and separated DNA fragments were visualized by staining with 0.5  $\mu$ g/ml EtBr and viewing under UV light.

5 In some cases, PCR generated fragments were sized by capillary electrophoresis using an ABI 310 Prism genetic analysis system (Applied Biosystems, Foster City, CA USA). The forward primer of the WMS44 primer pair was synthesized with a 5'-FAM label to generate fragments for this analysis. Samples were injected using a 5 second  
10 duration into a 20 cm capillary and separated under standard running conditions. Fragments were sized using the local Southern sizing method provided with the instrument.

3. RAPD analysis of bush rat DNA3.1 Template

15 Preparation of genomic DNA from seven individual bush rats (*Rattus fuscipes greyii*) was performed as described in Campbell *et al.*, 1995, Mol. Ecol. 4 407 which is herein incorporated by reference.

3.2 Primer

20 The primer used was primer A9 (Operon Technologies, Alameda CA) having the nucleotide sequence according to SEQ ID NO:5.

3.3. Thermal cycling

RAPD reactions were performed in 10  $\mu$ l volumes containing:-

25 1 ng/ $\mu$ l template;  
400 nM primer;  
250  $\mu$ g/ml BSA;  
1 Unit Taq polymerase;  
1 x 30 mM MgCl<sub>2</sub> reaction buffer (Idaho Technologies);  
30 50 mM each dNTP; and  
1:25000 dilution SYBR Green I.

Cycling conditions for RAPD reactions are as per Skroch &

Nienhuis, 1992, *The Rapid Cyclist 1* 9-10, which is herein incorporated by reference. Briefly, amplification is divided into two steps:-

- (i) 1 minute at 92°C; 7 seconds at 42 °C and 70 seconds at 72°C for 2 cycles; and
- 5 (ii) 1 second at 92°C; 7 seconds at 42°C and 70 seconds at 72°C for 38 cycles, followed by 4 minutes at 72°C.

**3.4 RAPD product analysis**

10 Amplification products were analyzed by electrophoresis on a 1% agarose gel. Separated DNA fragments were visualized by staining with EtBr and viewing under UV light.

**4. Melt curve analysis**

15 DNA melt curve analysis was performed using a *LightCycler* apparatus (Idaho Technology, Idaho Falls, ID, USA). The *LightCycler* uses glass capillaries as a reaction vessel and cuvette allowing for rapid temperature modulation and homogeneous reaction conditions (Wittwer *et al.*, 1997, *Biotechnology* 22 176).

20 SYBR Green I was used as the reporter molecule, and 2 µl of a 1:50,000 dilution was added per 10 µl of PCR reaction. Maximal excitement of SYBR Green I occurs at 497 nm with a secondary excitation peak at 254 nm; the fluorescence emission peak is centred at 520 nm.

25 Fluorescence emission by SYBR Green I was measured during a temperature increase from 60°C to 95°C at a transition rate of 0.1°C per second, with a fluorescent acquisition duration of 20 milliseconds for the Rrn5 measurements, and 3 milliseconds for the WMS44 measurements.

The conditions applicable to fluorescence monitoring of RAPD samples involved a 1:25000 dilution of SYBR Green I in each sample together with a temperature increase from 50°C to 92°C at the rate of 0.1°C per second.

30 Fluorescence emission (*F*) data and temperature (*T*) was converted to a melt curve (*F* vs *T*), and the negative derivative of the melt

curve plotted against temperature ( $-dF/dT$  vs  $T$ ).  $T_m$  values were obtained from the peak of the  $-dF/dT$  vs  $T$  curve. All melt curves, and information derived therefrom, were generated by software accompanying the *Light Cycler*. Post-hoc analyses of  $T_m$  values and the differences therebetween 5 were performed using the *Statistica V4 for Windows* program (Statsoft, Tulsa, OK, USA).

All melt curve analyses were performed using samples obtained from "end-point" PCR reactions. All analyses were replicated three times using separate reaction mixes.

10 **RESULTS AND DISCUSSION**

Ribosomal genes are commonly used in genotyping, particularly in plants. The identification of cereal plant varieties based on polymorphic 5S ribosomal (Rrn5) gene sequences using gel electrophoresis and hybridization has been previously described (Ko *et al.*, 1994, *J. Cereal Res.* 19 101). The Rrn5 gene family exists as tandem repeats of well conserved units of 120 base pairs each, each unit separated by a variable spacer region. The spacer may differ in length and in the number of copies of repeat units, this variation occurring between and within species. Species-specific PCR-amplification products 20 were obtained by using consensus primers based on the conserved 5S coding region.

Amplification products from each species tested are shown in FIG. 1. Rye produced a major ~400 bp product, however each of the other species produced an array of products, with no particular major product. Rice, wheat and maize were most similar in respect of the 25 amplified products resolved by gel electrophoresis, which was reflected by melt curve analysis. Melt curves were used to derive the  $-dF/dT$  vs  $T$  curves shown in FIG. 2, and  $T_m$  values calculated from the curves in FIG. 2 are shown in TABLE 1.

30 Visual inspection and comparison of each curve in FIG. 2 provided qualitative evidence that barley was distinct from all other

species. Also, the double peak evident for maize and rye (FIG. 2) distinguished these species from others.

The  $T_m$  values shown in TABLE 1 provided another type of information obtainable from the curves in FIG. 2. Each melt curve analysis 5 was performed three times,  $-dF/dT$  vs  $T$  curves were derived from each (FIG. 2), and mean  $T_m$  values together with standard deviations derived therefrom.

The statistical significance of the  $T_m$  differences is expressed as a Least Significance Difference using letter codes. 10 According to this code,  $T_m$  values that are not significantly different are indicated by shared letters, whereas  $T_m$  values which are significantly different share no letters.

With reference to the data in TABLE 1, barley was clearly distinguishable from each of the other species, and rye was significantly 15 different to oats on the basis of  $T_m$ . However, the distinction between maize and wheat or rice evident from FIG. 2 was not reflected by comparison of  $T_m$  values. Thus, this is an example where a qualitative inspection of respective  $-dF/dT$  v  $T$  curves can provide useful genotyping information not provided by a comparison of  $T_m$  values.

20 Based on these data, genotyping by melt curve analysis was capable of distinguishing between a variety of related cereal species. However, subsequent experiments also revealed this method of genotyping to be capable of distinguishing at the sub-species level.

25 Microsatellite or simple sequence repeats (SSR) are highly variable and ubiquitously expressed non-coding regions interspersed throughout eukaryotic genomes. PCR primers based on conserved regions of genomic DNA flanking microsatellites therefore provide a means of amplifying microsatellite regions from a wide variety of plants and animals. Microsatellite analysis is particularly valuable for genotypic 30 organisms where genetic variation is low. Furthermore, due to the extreme variability of microsatellite sequences, differences in

microsatellite PCR products can be used to distinguish between groups of individuals within a species. The present inventors have used the WMS44 microsatellite sequence found in the genome of wheat to discriminate between five different wheat cultivars and two wheat/rye hybrids (*Triticale* varieties) as a demonstration of sub-species genotyping by melt curve analysis.

Analysis of PCR-amplification products by PAGE revealed a typical array of fragments ranging in size from approximately 92 bp to 445 bp (FIG. 3), probably due to the presence of multiple priming sites in the complex hexaploid genome of wheat (Roder *et al.*, 1995, *supra*). Each variety tested, except the non-hexaploid *Triticale* varieties *Tiga* and *Tahara*, exhibited a ~180 bp fragment, as well as a number of other variably sized bands. Each ~180 bp fragment was purified and sized by capillary electrophoresis, the exact size of each being listed in TABLE 2.

The *Chinese Spring* 182 bp fragment was used as a reference, and sequencing of this 182 bp fragment revealed a "stutter peak" typical of microsatellites amplified from wheat (Plashke *et al.*, 1995, *Theor. Appl. Gen.* 91 1001). Capillary electrophoresis is an example of a state-of-the-art electrophoretic method used in genotyping, however, the method requires highly specialized equipment, is technically difficult and is time consuming. This method provided useful genotypic information in respect of five of the seven samples tested, and was able to distinguish four different genotypes.

The array of WMS44 microsatellite amplification products shown in FIG. 3 were subjected to melt curve analysis for all varieties except *Chinese Spring*. The  $-dF/dt$  v  $T$  curves are shown in FIG. 4, and the  $T_m$  values derived therefrom are shown in TABLE 3.

It was evident from these results that melt curve analysis was powerful enough to distinguish between:-

(i) cultivars sharing limited genetic relatedness, such as the wheat/rye hybrids *Tiga* or *Tahara* and each of the

wheat cultivars; and

(ii) genotypes at the sub-species level, for example, the wheat cultivars *Molineux* and *Gamenya* or *Halbert*.

In this regard, melt curve analysis was comparable to state 5 of the art electrophoretic methods such as capillary electrophoresis. However, the advantage of melt curve analysis is that it can be performed more quickly and with considerably less technical difficulty.

The present inventors have shown from this example that 10 the smallest distinguishable difference based on  $T_m$  alone was 0.8°C, which corresponded to a 2% change in G + C content for a given sequence length and salt concentration. Improvements in reporter molecules, instrumentation and analysis methods should even further improve this high level of resolution in the future.

The experiments using RAPD as a means of randomly 15 amplifying genomic DNA further demonstrate the power of the method of the invention. The RAPD fragments shown in FIG. 5 were generated using a PCR primer of arbitrary sequence, without regard to any known or predictable sequence polymorphism(s) in the seven individual bush rats tested. As can be seen clearly in FIG. 5, multiple RAPD products were 20 generated.

Six of the seven samples (corresponding to rats #3-8) were then subjected to melt curve analysis in triplicate, together with a negative control containing all reaction components except template.

Each replicate analysis yielded a single  $-dF/dT$  vs  $T$  melt 25 curve, with none of the melt curves having more than one resolvable peak. The data in TABLE 4 show the  $T_m$  values derived from each melt curve, together with a mean and standard deviation for respective samples corresponding to each individual rat tested. One way analysis of variance indicated significant differences in  $T_m$  estimates attributable to 30 rat genotype differences (significant in cases where  $p < 0.0500$ ; actual value of  $p = 0.005342$ ) and least significant difference testing indicated

which pairwise differences were significant (TABLE 5).

To test the possibility that fluctuations in the concentration of template present in RAPD reactions led to the observed  $T_m$  differences, the template concentration was varied over a 100-fold range.  $T_m$  estimates ranged from 84.30°C to 84.79°C with a mean and standard deviation of  $84.45^{\circ}\text{C} \pm 0.20^{\circ}\text{C}$ . The uniformity of measured  $T_m$  over such a wide template concentration range supported the conclusion that differences in derived  $T_m$  values reflected nucleotide sequence differences between RAPD samples, and were not merely due to quantitative variations between samples.

A key and unexpected finding from all of the experiments disclosed herein was that although there were multiple amplification products present in all samples (see FIGS. 1, 3 and 5), these multiple products produced a single  $-dF/dT$  vs  $T$  melt curve having a single peak for each sample except the cereal species maize and rye (FIG. 2), which each provided a melt curve having a pair of resolvable maxima or peaks. It must be noted that overall, there was no correlation between the number of resolvable peaks and the number of PCR-generated DNA fragments present in a sample.

In the case of maize and rye, the peak which occurred at the higher temperature was chosen for derivation of a  $T_m$  value, as it was at this higher temperature that the melting transition occurred for the majority of the DNA in these samples. It seems that some DNA fragments in the maize and rye samples were sufficiently different in their properties to show two resolvable maxima. Maize DNA fragments probably shared greater sequence homology and GC content than did rye DNA, as their major melt curve peaks were similar despite differences in DNA fragment sizes (FIG. 1). In contrast, the PCR products in oat, rice, wheat and barley samples produced melt curves with single peaks, probably because the PCR products shared similar nucleotide sequence and GC content.

In all cases, a single  $T_m$  value was calculable, which enabled genetic relationships to be determined between cereal species (TABLE 1), wheat cultivars and wheat/rye hybrids (TABLE 3) and individual bush rats (TABLE 4).

5 This result was not expected in light of Ririe *et al.*, 1997, *supra*, where a mixed sample containing distinct PCR products was resolved into distinct  $-dF/dT$  vs  $T$  peaks and  $T_m$  values corresponding to each PCR product. The results of Ririe *et al.*, 1997, *supra* had suggested that melt curve analysis would be restricted to comparing a single DNA  
10 fragment with another DNA fragment, where such fragments were substantially different in terms of nucleotide sequence. It was contended that the limit of this technology would be at the level of comparing heteroduplex DNA with homoduplex DNA, each corresponding to allelic forms of a particular polymorphic locus (Ririe *et al.*, 1997, *supra*).  
15 Accordingly, all of the prior art experimental evidence supporting melt curve analysis in genotyping has focussed on the discrimination of allelic sequences of one or a few nucleotides, this discrimination relying on detecting base-pair mismatches between differing sequences (see in particular, WO97/46714). This is particularly evident where discrimination  
20 is achieved by identifying distinct homoduplex and heteroduplex DNA-derived melt curves in a sample obtained from a heterozygote, or detecting a single homoduplex DNA-derived melt curve in a sample obtained from a homozygote. Such an approach has not lent itself to melt curve analysis of samples containing polynucleotide sequences having  
25 unknown potential for exhibiting sequence polymorphisms, even more particularly when the samples contain multiple different polynucleotide sequences.

In contrast, the present inventors have shown that melt curve analysis can be applied to the comparison of samples comprising  
30 multiple polynucleotide sequences, such as generated by PCR amplification of microsatellites, RAPD, AFLP or even RFLP. In many

cases, the precise nature of the sequence polymorphisms detectable by the method of the invention are unknown. Such a method would be impossible if genotypic differences were restricted to those known to be manifestable as a base-pair mismatch detectable by melt curve analysis.

5 Major attractions of genotyping by the method of the invention are its simplicity and speed, analyses being complete within minutes of obtaining the samples for melt curve analysis. Other detection technologies such as fluorescence energy transfer (see for example WO 97/46714), molecular beacons (Tyagi & Kramer, 1996, *Nature Biotechnology* **14** 303), or fluorescent probe hybridization (Wittwer *et al.*, 1997, In: *Gene Quantification*, Ed. Ferre, F, Birkhauser, NY; and also WO97/46714) are also sensitive, and in some cases can provide quantitative results. However, all are demanding in terms of complexity, time, cost and portability.

10 15 The method of this invention is applicable to the study of any entities that possess a genome. Examples of these genotyping applications could be in the area of inherited human disease, detecting pathogens, viral typing and identification, genotyping for organ transplantation, food analysis and selective breeding of plants and 20 animals. The simplicity of the technology, and the relatively compact apparatus used for melt curve analysis (together with the potential for miniaturization), make it possible that the method of genotypic analysis could be used "in the field". This portability would be particularly useful 25 for genotyping crop species and commercially bred animals in remote areas away from the sophisticated laboratory facilities normally required for genotyping.

TABLES**TABLE 1** Genotypic comparisons of cereal plant species according to  $T_m$ 

Species	$^1T_m$ ("C) $\pm$ SD	$^2LSD$
Barley	83.4 $\pm$ 1.3	
Oat	85.5 $\pm$ 0.2	a
Maize	86.4 $\pm$ 0.1	ab
Wheat	86.5 $\pm$ 0.4	abc
Rice	86.5 $\pm$ 0.2	abc
Rye	86.9 $\pm$ 0.0	bc

**TABLE 2** Size comparison of the major WMS44 microsatellite amplification product in five wheat strains and two wheat/rye hybrids

CULTIVAR <sup>1</sup>	FRAGMENT SIZE (BASE PAIRS)
<i>Tiga</i>	absent
<i>Tahara</i>	absent
<i>Gamenya</i>	172
<i>Halbert</i>	172
<i>Cunningham</i>	186
<i>Molineux</i>	186
<i>Chinese Spring</i>	182

**TABLE 3** Genotypic comparison of four wheat cultivars and two wheat/rye hybrid cultivars according to  $T_m$ 

CULTIVAR <sup>1</sup>	<sup>1</sup> $T_m$ (°C) ± SD	LSD <sup>2</sup>
<i>Tahara</i>	82.4 ± 0.5	a
<i>Tiga</i>	82.8 ± 0.3	ab
<i>Gamenya</i>	83.3 ± 0.4	bc
<i>Halbert</i>	83.5 ± 0.1	bcd
<i>Cunningham</i>	83.8 ± 0.5	cde
<i>Molineux</i>	84.3 ± 0.7	e

**TABLE 4** Summarized  $T_m$  estimates for RAPD samples.

SAMPLE	Expt 1	Expt 2	Expt 2	$T_m$ (°C) ± SD
<i>Rat #3</i>	83.47	84.48	84.28	84.07 ± 0.53
<i>Rat #4</i>	83.14	84.13	83.94	83.73 ± 0.52
<i>Rat #5</i>	82.84	82.70	83.21	83.25 ± 0.83
<i>Rat #6</i>	83.63	83.49	84.00	83.71 ± 0.26
<i>Rat #7</i>	84.46	84.27	84.39	84.37 ± 0.10
<i>Rat #8</i>	85.26	85.08	85.19	85.18 ± 0.09
<i>Negative control</i>	82.62	82.80	82.95	

**TABLE 5** Pairwise comparisons of individual bush rats according to  
<sup>1</sup>LSD testing of  $T_m$  values <sup>2</sup>

	Rat#4	Rat#5	Rat#6	Rat#7	Rat#8
Rat #3	0.39641	0.05375	0.35741	0.45768	0.01475
Rat #4		0.23202	0.93942	0.12550	0.00290
Rat #5			0.26038	0.01319	0.00033
Rat #6				0.11026	0.00252
Rat #7					0.05984

LEGENDS**TABLE 1**

<sup>1</sup> Average of three replicates  
<sup>2</sup> Species sharing the same letter are not significantly different at  
 5  $p=0.05$

**TABLE 2**

<sup>1</sup> *Tiga* and *Tahara* are wheat/rye hybrids

**TABLE 3**

<sup>1</sup> Average of three replicates  
 10 <sup>2</sup> Cultivars sharing the same letter are not significantly different at  
 $p=0.05$   
<sup>3</sup> *Tiga* and *Tahara* are wheat/rye hybrids.

**TABLE 5**

<sup>1</sup> Least Significant Differences test where statistically significant  
 15 differences between individual rats are evident when  $p < 0.0500$ .  
<sup>2</sup>  $T_m$  values used in calculating  $p$  were identical to the mean  $T_m$   
 values shown in TABLE 4.

**FIG. 1**

PCR amplified products from six cereal plant species. Sequences were  
 20 amplified using consensus primers to conserved regions of the 5S  
 Ribosomal RNA locus in cereal plants. 10  $\mu$ l of each sample was  
 subjected to 2% agarose gel electrophoresis and visualized under UV  
 light after EtBr staining. O = oat; R = rice; M = maize; W = wheat; Ry =  
 25 rye; B = barley. Std = molecular size standards consisting of a 50 bp  
 ladder. Major 500 bp and 250 bp standards are indicated.

**FIG. 2**

$-dF/dT$  vs  $T$  curves derived from the melt curves for each sample shown  
 in FIG. 1. Fluorescence acquisition duration was 20 milliseconds with a  
 temperature ramp from 60-95°C at 0.1°C per second.  $-dF/dT$  vs  $T$  curves  
 30 were derived from melt curves using a sample loading of eight, points to  
 average setting of 15 and no baseline subtraction.

**FIG. 3**

Variability in PCR fragments amplified from WMS44 microsatellites from five wheat cultivars and two wheat/rye hybrids. 10  $\mu$ l of each PCR reaction was subjected to 10% PAGE and visualized under UV light after staining with EtBr. M = molecular size standards consisting of a 50 bp ladder; 50 bp, 250 bp and 500 bp standards are indicated. Gam = *Gamenya*; Hal = *Halbert*; Cun = *Cunningham*; Tig = *Tigara*; Tah = *Tahara*, Mol = *Molineux*. Tig and Tah are the wheat/rye hybrids.

**FIG. 4**

- $dF/dT$  vs  $T$  melt curves for the PCR fragments amplified from WMS44 microsatellites from the four wheat cultivars and the two wheat/rye hybrids in FIG. 3. Fluorescence acquisition duration was 3 milliseconds with a temperature ramp from 60-95°C at 0.1°C per second. - $dF/dT$  vs  $T$  melt curves were derived from melt curves using a sample loading of seven, 15 points to average setting of 20 and no baseline subtraction.

**FIG. 5**

Agarose gel electrophoresis analysis of RAPD products for each of seven bush rats. The lanes from left to right are as follows: (1) 100 bp ladder; (2) rat #3; (3) rat # 4; (4) rat #5; (5) rat #6; (6) rat #7; (7) rat #8; and (8) rat 20 #9.

## SEQUENCE LISTING

<110> Southern Cross University  
Grains Research and Development Corporation

<120> A Method of Genotyping

<130> henry-shep

<140>  
<141>

<150> PO 8824

<151> 1997-08-29

<160> 5

<170> PatentIn Ver. 2.0

<210> 1  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Non  
species-specific primer for amplifying cereal Rrn5  
ribosomal locus

<400> 1

tgggaagtcc tcgtgttgca

20

<210> 2

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Non  
species-specific primer for amplifying cereal Rrn5  
ribosomal locus

<400> 2

ttagtgctg gtatgatcgc

20

<210> 3

<211> 19

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer for  
amplifying wheat WMS44 microsatellite locus

35

<400> 3  
gttgagctt cagttcgcc 19

<210> 4  
<211> 19  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer for  
amplifying wheat WMS44 microsatellite locus

<400> 4  
actggcatcc actgagctg 19

<210> 5  
<211> 10  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: A9 Primer  
(Operon Technologies)

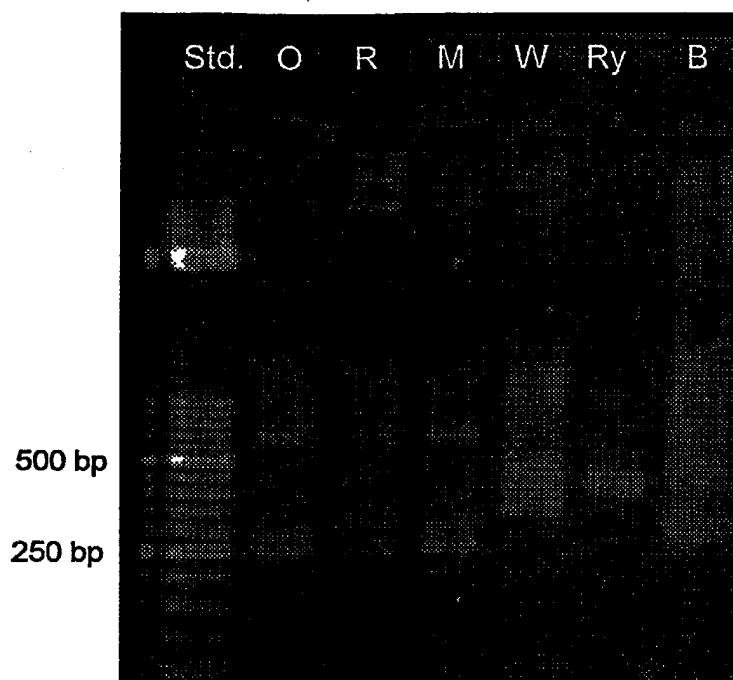
<400> 5  
gggttaacgcc 10

CLAIMS

1. A method of genotyping including the steps of:-
  - (i) producing two or more samples, each of which includes two or more double-stranded polynucleotide sequences representative of a genotype of an individual or group of individuals;
  - (ii) adding a fluorescent reporter molecule to the samples produced in step (i)
  - (iii) modulating temperature whereby a conformation of said two or more double-stranded polynucleotide sequences is altered, and measuring a signal emanating from said reporter molecule, which signal changes as the conformation of said one or more double-stranded polynucleotide sequences is altered;
  - (iv) using the signal measured in (iii) to construct a single  $-dF/dT$  melt curve corresponding to the two or more polynucleotide sequences; and
  - (v) comparing respective single  $-dF/dT$  melt curves, or single  $T_m$  values derived from each melt curve, in order to identify a genetic relationship between each said individual or group of individuals.
2. The method of Claim 1 wherein respective single  $T_m$  values are compared at step (v) in order to identify a genetic relationship between each said individual or group of individuals.
3. The method of Claim 1 or Claim 2, wherein the polynucleotide sequences indicative of genotype are produced by polynucleotide sequence amplification techniques.
4. The method of Claim 3, wherein the polynucleotide sequence amplification technique is polymerase chain reaction (PCR).
5. The method of Claim 4, wherein PCR involves Random Amplification of Polymorphic DNA (RAPD).

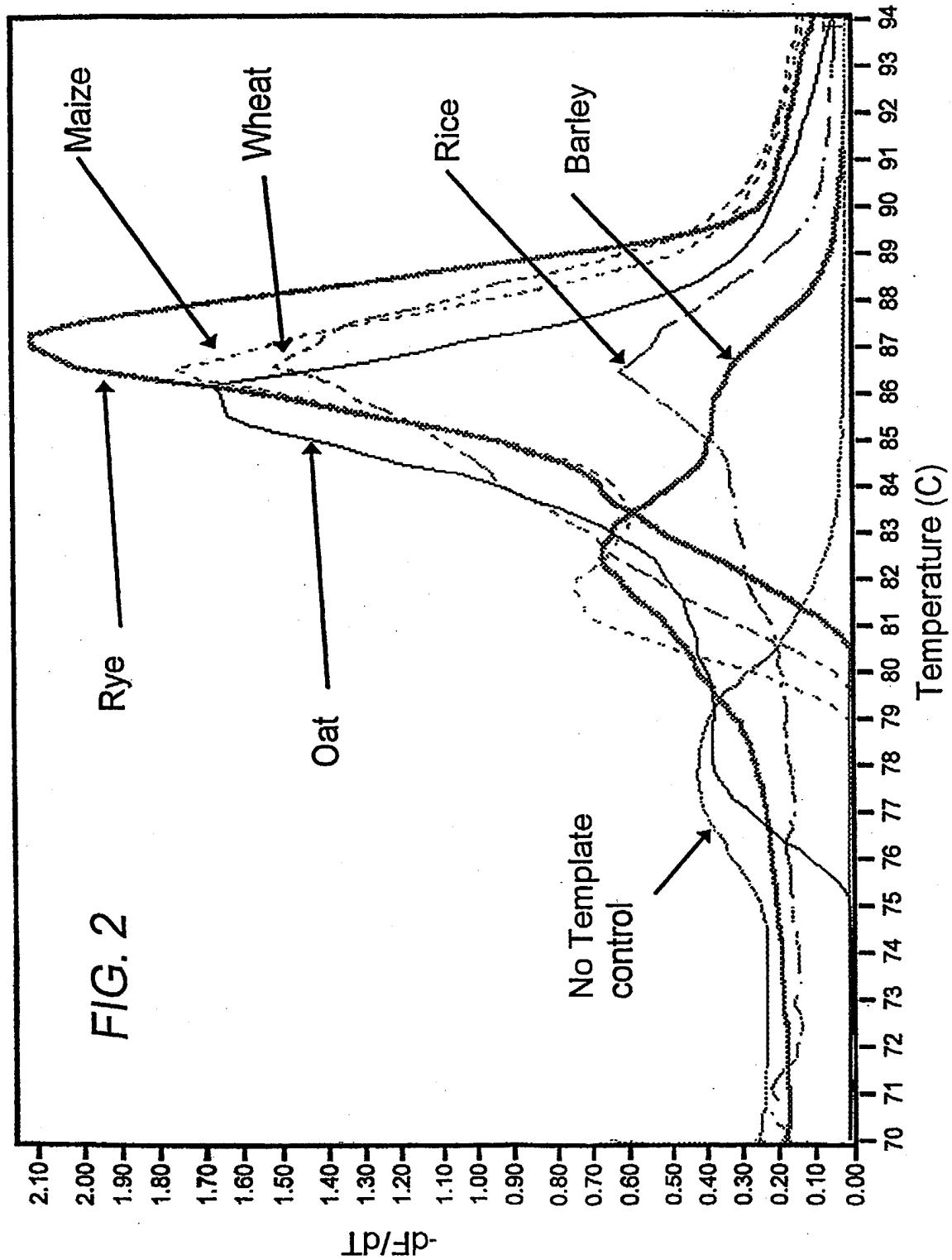
6. The method of any one of Claims 1-5, wherein the polynucleotide sequences indicative of genotype have repetitive sequence elements.
7. The method of Claim 6, wherein the polynucleotide sequences indicative of genotype have at least one microsatellite repeat sequence.
8. The method of Claim 7, wherein the microsatellite repeat sequence corresponds to the WMS44 microsatellite locus.
9. The method of Claim 6, wherein the polynucleotide sequences indicative of genotype have at least one ribosomal repeat sequence.
10. The method of Claim 9, wherein the ribosomal repeat sequence corresponds to the Rrn5 ribosomal repeat.

1/5



**FIG. 1**

2/5



3/5

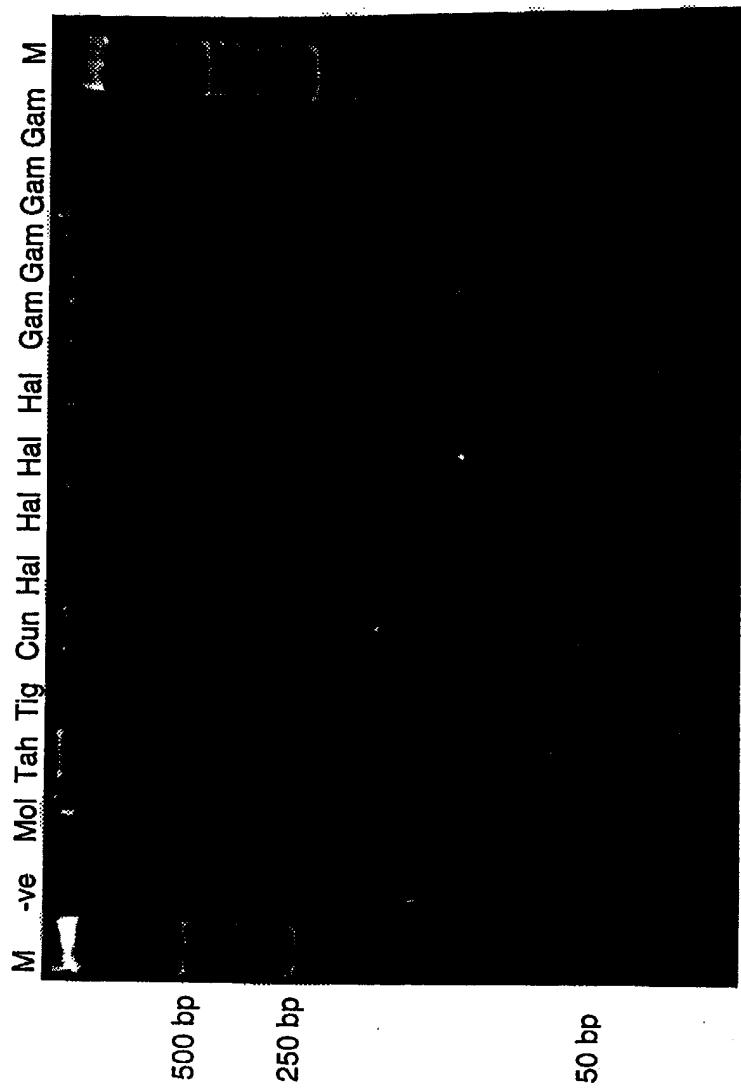


FIG. 3

4/5

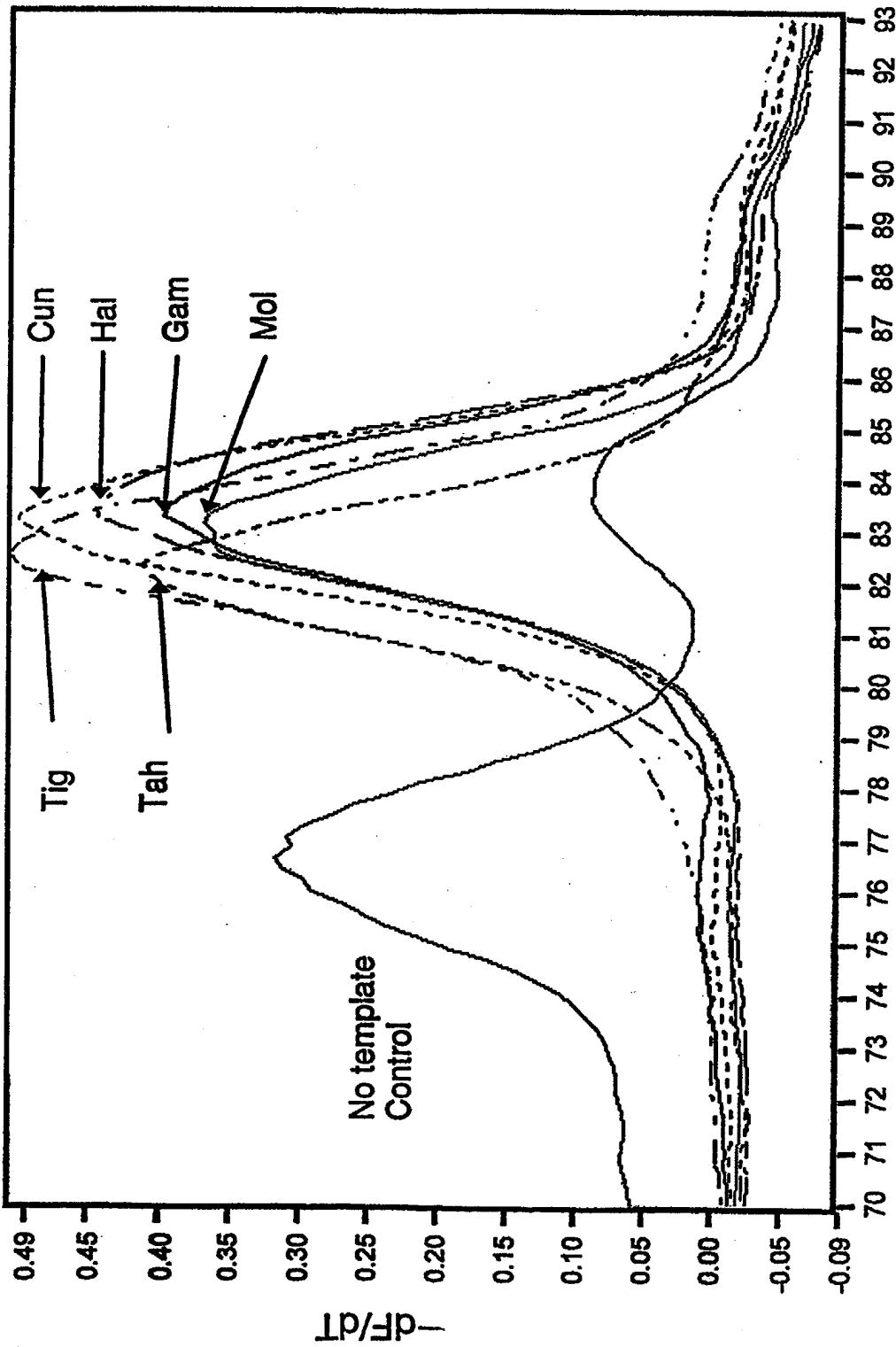


FIG. 4  
Temperature (C)

5/5

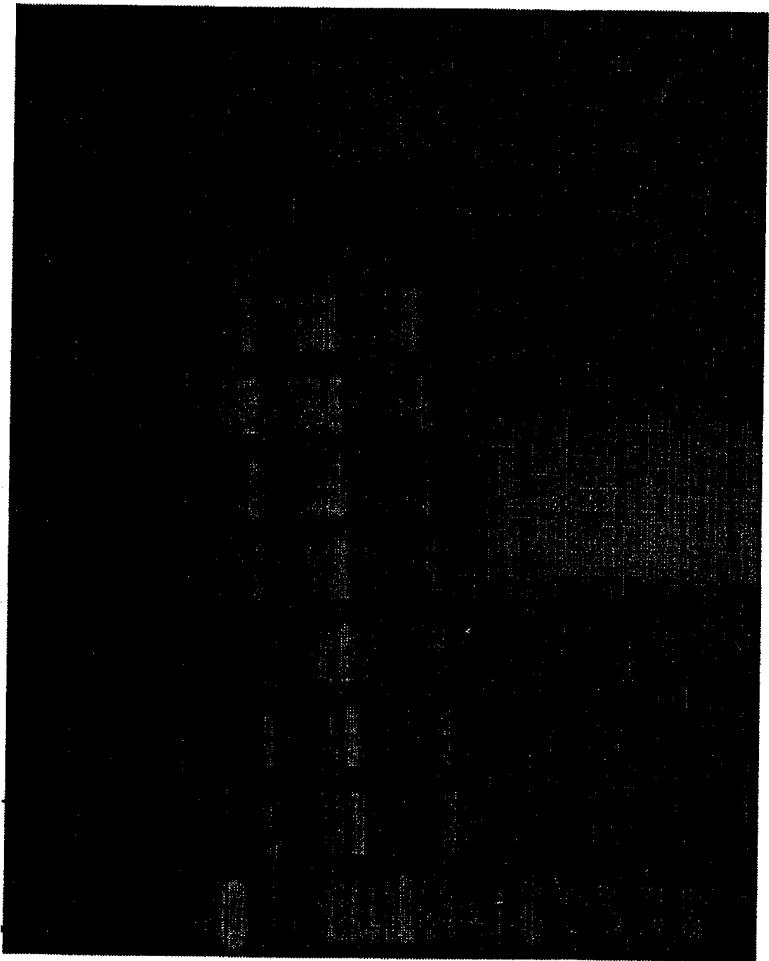


FIG. 5

# INTERNATIONAL SEARCH REPORT

International Application No.  
PCT/AU 98/00700

<b>A. CLASSIFICATION OF SUBJECT MATTER</b>		
Int Cl <sup>6</sup> : G01N 33/483; C12Q 1/68		
According to International Patent Classification (IPC) or to both national classification and IPC		
<b>B. FIELDS SEARCHED</b>		
Minimum documentation searched (classification system followed by classification symbols)		
IPC	G01N, C12Q	
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
AU:	IPC as above	
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)		
WPAT	: fluoresc: and melt: and (temp: or curve? or profile?) and (DNA or double(stranded))	
STN (Chemical Abstracts)	: melt? and, (curve? or profile?) and fluorescenc	
Medline	: mell and (curve? or profile: or temp:) and fluoresc? and DNA/CT	
<b>C. DOCUMENTS CONSIDERED TO BE RELEVANT</b>		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	PCR Methods and Applications volume 4 no. 6 (June 1995) Eggerding, F A "A One-step Coupled Amplification and Oligonucleotide Ligation Procedure for Multiplex Genetic Typing" pages 337-345 pages 338 and 341, Table 1	1-10
X	Cancer Research volume 37 no. 4 (April 1977) Melamed, et al "Nucleic Acid Content and Nuclear Chromatin Structure of Human Bladder Cell Culture Lines as Studied by Flow Cytofluorometry" pages 1254-1231 entire document	1-10
<input checked="" type="checkbox"/> Further documents are listed in the continuation of Box C		<input type="checkbox"/> See patent family annex
<p>* Special categories of cited documents:</p> <p>"A" document defining the general state of the art which is not considered to be of particular relevance</p> <p>"E" earlier document but published on or after the international filing date</p> <p>"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)</p> <p>"O" document referring to an oral disclosure, use, exhibition or other means</p> <p>"P" document published prior to the international filing date but later than the priority date claimed</p> <p>"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention</p> <p>"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone</p> <p>"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art</p> <p>"&amp;" document member of the same patent family</p>		
Date of the actual completion of the international search 21 September 1998	Date of mailing of the international search report 28 SEP 1998	
Name and mailing address of the ISA/AU AUSTRALIAN PATENT OFFICE PO BOX 200 WODEN ACT 2606 AUSTRALIA Facsimile No.: (02) 6285 3929	Authorized officer  <b>STEVEN CHEW</b> Telephone No.: (02) 6283 2248	

**INTERNATIONAL SEARCH REPORT**

International Application No.

PCT/AU 98/00700

<b>C (Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT</b>		
<b>Category*</b>	<b>Citation of document, with indication, where appropriate, of the relevant passages</b>	<b>Relevant to claim No.</b>
X	<u>International Journal of Systematic Bacteriology</u> volume 46 no. 4 (October 1996) Elomari, et al "DNA Relatedness among <u>Pseudomonas</u> Strains Isolated from Natural Mineral Waters and Proposal of <u>Pseudomonas veronii</u> sp. nov." pages 1138-1144 Table 1 page 1139; "Results" page 1141	1-10
X	<u>Chromosoma</u> volume 95 (1987) Kusewitt, et al "Evolutionary Relationships of the Chinese hamster X Chromosome and Autosomes: A comparison using solution hybridisation techniques" pages 189-196 entire document	1-10
X	<u>Analytical Biochemistry</u> 245 (15 February 1997) Ririe et al "Product Differentiation by Analysis of DNA Melting Curves during the Polymerase Chain Reaction" pages 154-160 entire document	1-10